Computational Medicinal Chemistry in Drug Discovery

Schultes, S.

2013

document version
Publisher's PDF, also known as Version of record

Summary

The primary aim of the research described in this thesis, as defined in **Chapter 2**, was to show how computational chemistry methods can be developed and used to:

1) improve our understanding of physicochemical properties that are important in the development of biologically active molecules towards drug lead compounds.

2) improve our understanding of the chemical properties and molecular pharmacophore features that determine the biological activity of small molecules.

3) obtain insights in the molecular determinants and binding modes of small molecule ligands to pharmaceutically relevant protein targets.

My work particularly focused on: I) The definition of ligand efficiency scores as guidelines to control the size and lipophilicity of biologically active molecules during drug development, II) The identification of simple descriptors that are related to the melting point of drug-like compounds, III) The evaluation and combination of ligand-based virtual screening (LBVS) approaches for the identification of fragment-like bioactive molecules and IV) Structure-based modeling studies to elucidate ligand binding modes in the histamine H_4 receptor (H_4R).

Chapter 1 provides an overview of the drug research process. The development of potential drug candidates by medicinal chemists is explained and issues and challenges during this process are discussed. Drug candidates need to have good ADMET (absorption, distribution, metabolism, excretion and toxicology) properties and good affinity towards a defined drug target. Computational chemistry provides tools to support medicinal chemists in different stages (screening, hit to lead and lead optimization) of the drug finding process. Virtual screening supports hit finding, quantitative structure property relationship (QSPR) can be applied for ADMET optimization and modeling tools are useful for affinity optimization towards a drug target. As a drug target family of high interest, G Protein-Coupled Receptors (GPCRs) are introduced. A special focus has been put on H_4R as it is a promising new target in the GPCR field for modulating immune responses and inflammatory processes.

Chapter 2 defines the scope and the aim of the thesis.

Chapter 3 (**research aim 1)** reviews how ligand efficiency scores have been applied to guide the selection and optimization of fragment hits. Both molecular size and lipophilicity are important ADMET related parameters that need to be carefully controlled during the drug discovery process. Ligand efficiency scores, like ligand efficiency (LE)[1], binding

efficiency index (BEI)[2], ligand-lipophilicity efficiency (LLE)[3], ligand efficiency dependent lipophilicity (LELP)[4], fit quality (FQ)[5-7], size independent ligand efficiency(SILE)[8], that combine molecular size and lipophilicity have been proposed as a better guide during the drug discovery process than affinity on its own. Fragment-based screening (FBS) has become an established approach for hit identification. The advantage of fragments as starting points is that low molecular weight and hydrophilic compounds form the basis of the further optimization process, and the regular features to optimize affinity, such as molecular size and lipophilicity, are not initially constraining factors. Ligand efficiency scores are ideal parameters to guide the hit selection and optimization process.

In **Chapter 4** (**research aim 1**) the influence of simple molecular descriptors on the melting point temperature (T_m) has been investigated using a matched molecular pair (MMP) analysis. This method has been used to identify small structural differences between pairs of molecules, which are put into relation to changes in T_m. Our analysis shows that the number of hydrogen bond donors, hydrogen bond acceptors and rotatable bonds has a significant effect on the T_m of a molecule. Hydrogen bond donors have the most pronounced effect. Furthermore, the studies reveal that not ClogP but rather the number of bromine and iodine atoms has a marked effect on the T_m. The results of our MMP analysis are discussed within the context of drug solubility optimization.

In **Chapter 5** (**research aim 2**) the applicability of LBVS methods for fragments using different similarity methods and different consensus scoring approaches was tested. For this task, performance of all possible combinations of 14 similarity methods was tested. In addition, the dependency of the success of LBVS on similarity between the test and reference compounds was investigated. The results reveal that EDprints seems to be less dependent on similarity, and therefore is the most suitable method to enrich dissimilar scaffolds. Consensus scoring was shown to significantly increase enrichments of actives compared to the average enrichments. For group fusion (combining the similarity values obtained by comparing test compounds with several actives) ranked-by-vote and also (less frequently) the maximum similarity most often perform best. For data fusion (combining the scores of the different similarity methods) the mean-value and mean-rank most often perform best. Interestingly the best performing individual method is not necessarily included in the best combination of similarity methods and also combining the best individually performing methods does not necessarily lead to considerably higher enrichments than other method combinations. Finally the study shows that although LBVS, especially for small fragments, can be challenging, it can be applied for fragment compounds and is also suitable for enrichment of dissimilar scaffolds.

Chapter 6 (**research aim 3**) presents the combination of complementary *in silico* receptor modeling approaches with in vitro ligand structure-activity relationship (SAR) and protein site-directed mutagenesis studies to elucidate the binding modes of 2-aminopyrimidines and indolecarboxamides in the H_4R. In this study different ligand binding poses in combination with different protonation states in different H_4R modeling templates were investigated. Site directed mutagenesis studies revealed important interactions sites in the H_4R binding pocket. Using MD simulations these interactions could be studied in atomic detail, and formed the basis to explain ligand-specific mutation effects as well as subtle differences in SAR of the investigated H_4R ligand classes. The studies underline that a combined theoretical and experimental approach represents a powerful strategy to map ligand-protein interactions.

Chapter 7 (**research aim 3**) focuses on the elucidation of the binding conformation of 2-aminopyrimidine ligands with flexible side-chains in the H_4R binding pocket. In **Chapter 6** a binding mode for 2-aminopyrimidine ligands bearing a rigid phenyl moiety on position 6 of the pyrimidine was proposed. In this study 2-aminopyrimidine ligands with flexible side-chains were designed and synthesized to investigate the role of ligand conformation in H_4R ligand binding. By combining structure-activity relationships, protein-ligand modeling studies, and quantum mechanical calculations we have obtained new insights into the molecular determinants of H_4R-ligand binding and could elucidate ligand binding conformations and orientations in the H_4R binding pocket. Our combined ligand- and protein-ligand modeling method can be used as a general approach to investigate the binding modes of flexible side-chains of ligands in other protein targets.